```
Val Ala Leu Asp Tyr Ser Gln Ile Glu Leu Arg Val Leu Ala His Leu
    610             615             620

Ser Gly Asp Glu Asn Leu Ile Arg Val Phe Gln Glu Gly Lys Asp Ile
625             630             635             640

His Thr Gln Thr Ala Ser Trp Met Phe Gly Val Pro Pro Glu Ala Val
            645             650             655

Asp Pro Leu Met Arg Arg Ala Ala Lys Thr Val Asn Phe Gly Val Leu
            660             665             670

Tyr Gly Met Ser Ala His Arg Leu Ser Gln Glu Leu Ala Ile Pro Tyr
        675             680             685

Glu Glu Ala Val Ala Phe Ile Glu Arg Tyr Phe Gln Ser Phe Pro Lys
    690             695             700

Val Arg Ala Trp Ile Glu Lys Thr Leu Glu Glu Gly Arg Lys Arg Gly
705             710             715             720

Tyr Val Glu Thr Leu Phe Gly Arg Arg Arg Tyr Val Pro Asp Leu Asn
            725             730             735

Ala Arg Val Lys Ser Val Arg Glu Ala Ala Glu Arg Met Ala Phe Asn
            740             745             750

Met Pro Val Gln Gly Thr Ala Ala Asp Leu Met Lys Leu Ala Met Val
        755             760             765

Lys Leu Phe Pro Arg Leu Arg Glu Met Gly Ala Arg Met Leu Leu Gln
    770             775             780

Val His Asp Glu Leu Leu Leu Glu Ala Pro Gln Ala Arg Ala Glu Glu
785             790             795             800

Val Ala Ala Leu Ala Lys Glu Ala Met Glu Lys Ala Tyr Pro Leu Ala
            805             810             815

Val Pro Leu Glu Val Glu Val Gly Met Gly Glu Asp Trp Leu Ser Ala
        820             825             830

Lys Gly
```

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2502 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
ATGNNGGCGA TGCTTCCCCT CTTTGAGCCC AAAGGCCGGG TCCTCCTGGT GGACGGCCAC      60

CACCTGGCCT ACCGCACCTT CTTCGCCCTG AAGGGCCTCA CCACCAGCCG GGGCGAACCG     120

GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTGAAGGA GGACGGGGAC     180
```

```
NNGGCGGTGN TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA GGCCTACGAG    240

GCCTACAAGG CGGGCCGGGC CCCCACCCCG GAGGACTTTC CCCGGCAGCT CGCCCTCATC    300

AAGGAGCTGG TGGACCTCCT GGGGCTTGCG CGCCTCGAGG TCCCCGGCTA CGAGGCGGAC    360

GACGTNCTGG CCACCCTGGC CAAGAAGGCG GAAAAGGAGG GGTACGAGGT GCGCATCCTC    420

ACCGCCGACC GCGACCTCTA CCAGCTCCTT TCCGACCGCA TCGCCGTCCT CCACCCCGAG    480

GGGTACCTCA TCACCCCGGC GTGGCTTTGG GAGAAGTACG GCCTGAGGCC GGAGCAGTGG    540

GTGGACTACC GGGCCCTGGC GGGGGACCCC TCCGACAACC TCCCCGGGGT CAAGGGCATC    600

GGGGAGAAGA CCGCCCNGAA GCTCCTCNAG GAGTGGGGGA GCCTGGAAAA CCTCCTCAAG    660

AACCTGGACC GGGTGAAGCC CGCCNTCCGG GAGAAGATCC AGGCCCACAT GGANGACCTG    720

ANGCTCTCCT GGGAGCTNTC CCAGGTGCGC ACCGACCTGC CCCTGGAGGT GGACTTCGCC    780

AAGNGGCGGG AGCCCGACCG GGAGGGGCTT AGGGCCTTTC TGGAGAGGCT GGAGTTTGGC    840

AGCCTCCTCC ACGAGTTCGG CCTCCTGGAG GGCCCCAAGG CCCTGGAGGA GGCCCCCTGG    900

CCCCCGCCGG AAGGGGCCTT CGTGGGCTTT GTCCTTTCCC GCCCCGAGCC CATGTGGGCC    960

GAGCTTCTGG CCCTGGCCGC CGCCAGGGAG GGCCGGGTCC ACCGGGCACC AGACCCCTTT   1020

ANGGGCCTNA GGGACCTNAA GGAGGTGCGG GGNCTCCTCG CCAAGGACCT GGCCGTTTTG   1080

GCCCTGAGGG AGGGCCTNGA CCTCNTGCCC GGGGACGACC CCATGCTCCT CGCCTACCTC   1140

CTGGACCCCT CCAACACCAC CCCCGAGGGG GTGGCCCGGC GCTACGGGGG GGAGTGGACG   1200

GAGGANGCGG GGGAGCGGGC CCTCCTNTCC GAGAGGCTCT TCCNGAACCT NNNGCAGCGC   1260

CTTGAGGGGG AGGAGAGGCT CCTTTGGCTT TACCAGGAGG TGGAGAAGCC CCTTTCCCGG   1320

GTCCTGGCCC ACATGGAGGC CACGGGGGTN CGGCTGGACG TGGCCTACCT CCAGGCCCTN   1380

TCCCTGGAGG TGGCGGAGGA GATCCGCCGC CTCGAGGAGG AGGTCTTCCG CCTGGCCGGC   1440

CACCCCTTCA ACCTCAACTC CCGGGACCAG CTGGAAAGGG TGCTCTTTGA CGAGCTNGGG   1500

CTTCCCGCCA TCGGCAAGAC GGAGAAGACN GGCAAGCGCT CCACCAGCGC CGCCGTGCTG   1560

GAGGCCCTNC GNGAGGCCCA CCCCATCGTG GAGAAGATCC TGCAGTACCG GGAGCTCACC   1620

AAGCTCAAGA ACACCTACAT NGACCCCCTG CCNGNCCTCG TCCACCCCAG GACGGGCCGC   1680

CTCCACACCC GCTTCAACCA GACGGCCACG GCCACGGGCA GGCTTAGTAG CTCCGACCCC   1740

AACCTGCAGA ACATCCCCGT CCGCACCCCN CTGGGCCAGA GGATCCGCCG GGCCTTCGTG   1800

GCCGAGGAGG GNTGGGTGTT GGTGGCCCTG GACTATAGCC AGATAGAGCT CCGGGTCCTG   1860

GCCCACCTCT CCGGGGACGA GAACCTGATC CGGGTCTTCC AGGAGGGGAG GGACATCCAC   1920

ACCCAGACCG CCAGCTGGAT GTTCGGCGTC CCCCCGGAGG CCGTGGACCC CCTGATGCGC   1980

CGGGCGGCCA AGACCATCAA CTTCGGGGTC CTCTACGGCA TGTCCGCCCA CCGCCTCTCC   2040
```

```
CAGGAGCTTG CCATCCCCTA CGAGGAGGCG GTGGCCTTCA TTGAGCGCTA CTTCCAGAGC      2100

TTCCCCAAGG TGCGGGCCTG GATTGAGAAG ACCCTGGAGG AGGGCAGGAG GCGGGGGTAC      2160

GTGGAGACCC TCTTCGGCCG CCGGCGCTAC GTGCCCGACC TCAACGCCCG GGTGAAGAGC      2220

GTGCGGGAGG CGGCGGAGCG CATGGCCTTC AACATGCCCG TCCAGGGCAC CGCCGCCGAC      2280

CTCATGAAGC TGGCCATGGT GAAGCTCTTC CCCCGGCTNC AGGAAATGGG GGCCAGGATG      2340

CTCCTNCAGG TCCACGACGA GCTGGTCCTC GAGGCCCCCA AGAGCGGGC GGAGGNGGTG       2400

GCCGCTTTGG CCAAGGAGGT CATGGAGGGG GTCTATCCCC TGGCCGTGCC CCTGGAGGTG      2460

GAGGTGGGGA TGGGGGAGGA CTGGCTCTCC GCCAAGGAGT AG                        2502
```

(2) INFORMATION FOR SEQ ID NO:8:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 833 amino acids
          (B) TYPE: amino acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: unknown

      (ii) MOLECULE TYPE: peptide

      (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
Met Xaa Ala Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
1               5                   10                  15

Val Asp Gly His His Leu Ala Tyr Arg Thr Phe Phe Ala Leu Lys Gly
            20              25                  30

Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
        35                  40                  45

Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Xaa Val
    50              55                  60

Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Glu Ala
65              70                  75                  80

Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
                85                  90                  95

Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Xaa Arg Leu Glu
            100                 105                 110

Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Thr Leu Ala Lys Lys
            115                 120                 125

Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Arg Asp
    130                 135                 140

Leu Tyr Gln Leu Leu Ser Asp Arg Ile Ala Val Leu His Pro Glu Gly
145                 150                 155                 160

Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
                165                 170                 175
```